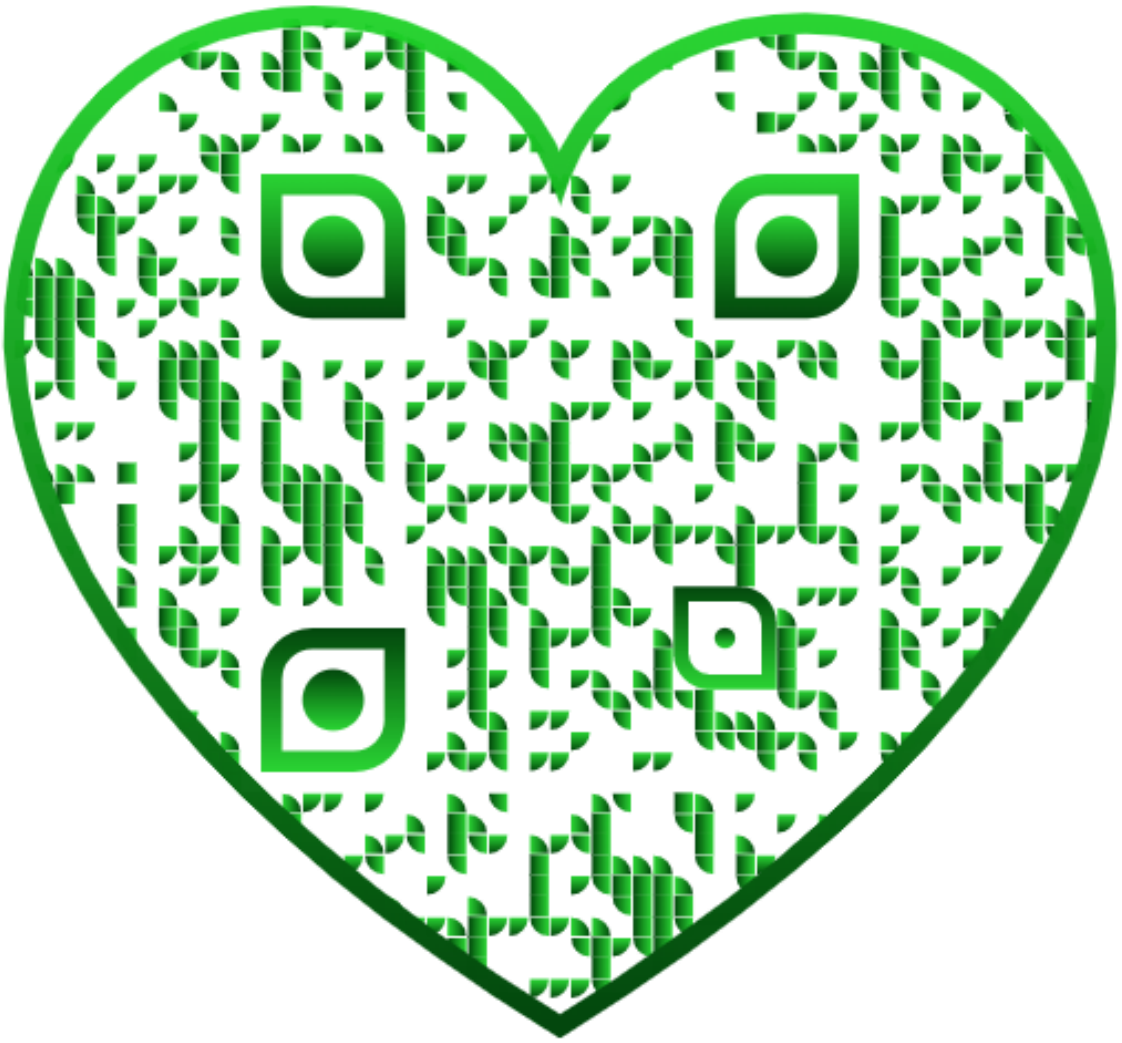
CH_8: Hypothesis testing

At the end of this chapter the student will be able to:

- Define the concepts of null and alternate hypothesis
- Differentiate between type I and type II errors
- Describe the different types of statistical tests used when samples are large and small
- Explain the meaning and application of P – values
- Define the concepts of degrees of freedom
- Explain measures of associations e.g. X^2 test



Definition

- **A statistical hypothesis is an <u>assumption</u> or a <u>statement</u>** which may or may not be true concerning one or more populations.
- The significance test or hypothesis test.
- A significance test enables us to measure the strength of the evidence which the data supply concerning some proposition of interest.

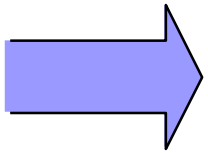
Hypothesis testing

- The **purpose** of hypothesis testing is to determine whether enough statistical evidence exists to enable us to conclude that a belief or hypothesis about a parameter is reasonable

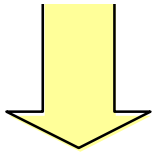
Hypothesis Testing Process

Assume the population mean age is 50.

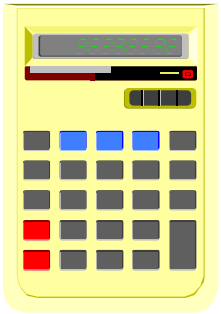
($H_0 : \mu = 50$)



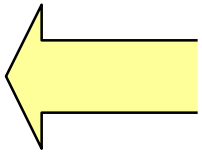
Identify the Population



Take a Sample



$(\overline{X} = 20)$



Is $\overline{X} = 20$ likely if $\mu = 50$?

No, not likely!

REJECT H_0

Type of statistical hypotheses

- There are two hypotheses:
  - ❖ Null hypotheses
  - ❖ Alternative hypotheses

1) Null hypotheses

- **Null hypothesis** – called the hypothesis of no difference or no association or no effect
- States that ‘’there’s no difference’’ between the hypothesized value and the population parameter value
- Is always about a **population parameter**, not about a sample

Null Hypothesis: H_0

- The null hypothesis (denoted by H_0) is a statement that the value of a population parameter (such as proportion, mean, or standard deviation) is equal to some claimed value.
- Always contains the "=" "≤" **or** "≥" sign
- We test the null hypothesis directly.
- Either reject H_0 or fail to reject H_0.

2) Alternative hypotheses

- Alternate to null hypothesis
- Says'' there's a difference between the hypothesized value and the population parameter value
- It is what we are trying to prove, i.e. the reason for the research question.

Alternative Hypothesis: H_1 or H_A

- The alternative hypothesis (denoted by H_1 or H_A) is the statement that the parameter has a value that somehow differs from the null hypothesis.
- The symbolic form of the alternative hypothesis must use one of these symbols: $\neq$, **< or >.**
- May or may not be accepted

Hypothesis

Example: Consider population mean

$H_0: \mu = \mu_0$
$H_A: \mu \neq \mu_0$

Two- tailed

$H_0: \mu \leq \mu_0$
$H_A: \mu > \mu_0$

One - tailed

$H_0: \mu \geq \mu_0$
$H_A: \mu < \mu_0$

One- tailed

Exercises

➔ ***State H_O and H_A for each of the following***

1) Is the average height of health science students 1.65 m or is it more?

2) Is the average height of MPH students 1.65 m or is it less?

3) Is the average height of MPH students 1.65 m or is it something different?

4) Are men and women infected with malaria in equal proportions, or is a higher proportion of men get malaria in Ethiopia?

Steps in hypothesis testing

1) State the statistical hypotheses

- There are two hypotheses:

  -Null hypotheses

  - Alternative hypotheses

2) Select the level of significance and determine critical value

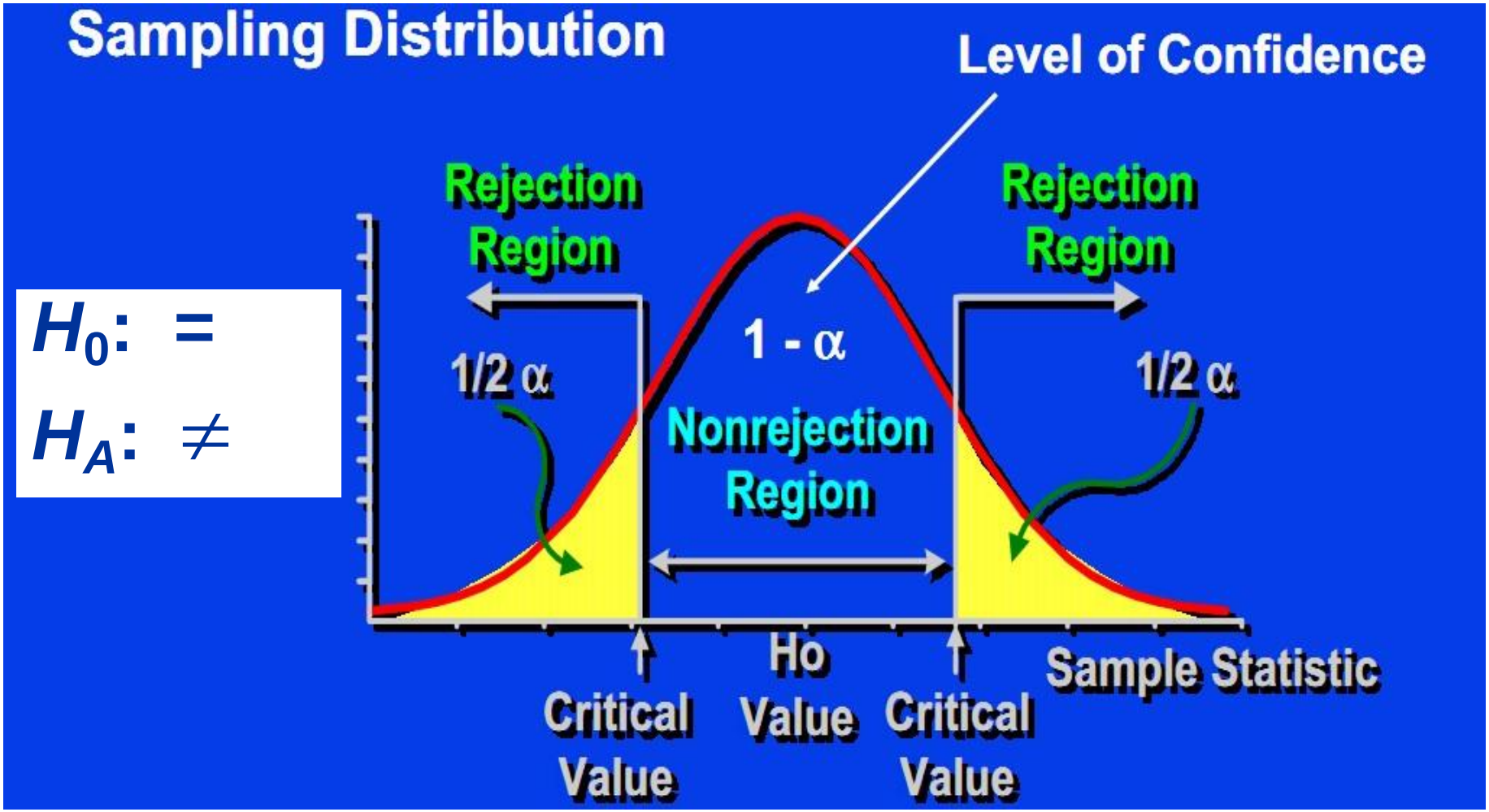
Level of Significance (α) : Defines rejection region $\mathbf{H_0}$ of the sampling distribution

- Called rejection region of sampling distribution
- Is designated by α , (level of significance)
  - Typical values are .01, **.05**, or .10
- Is selected by the researcher at the beginning
- Provides the critical value(s) of the test

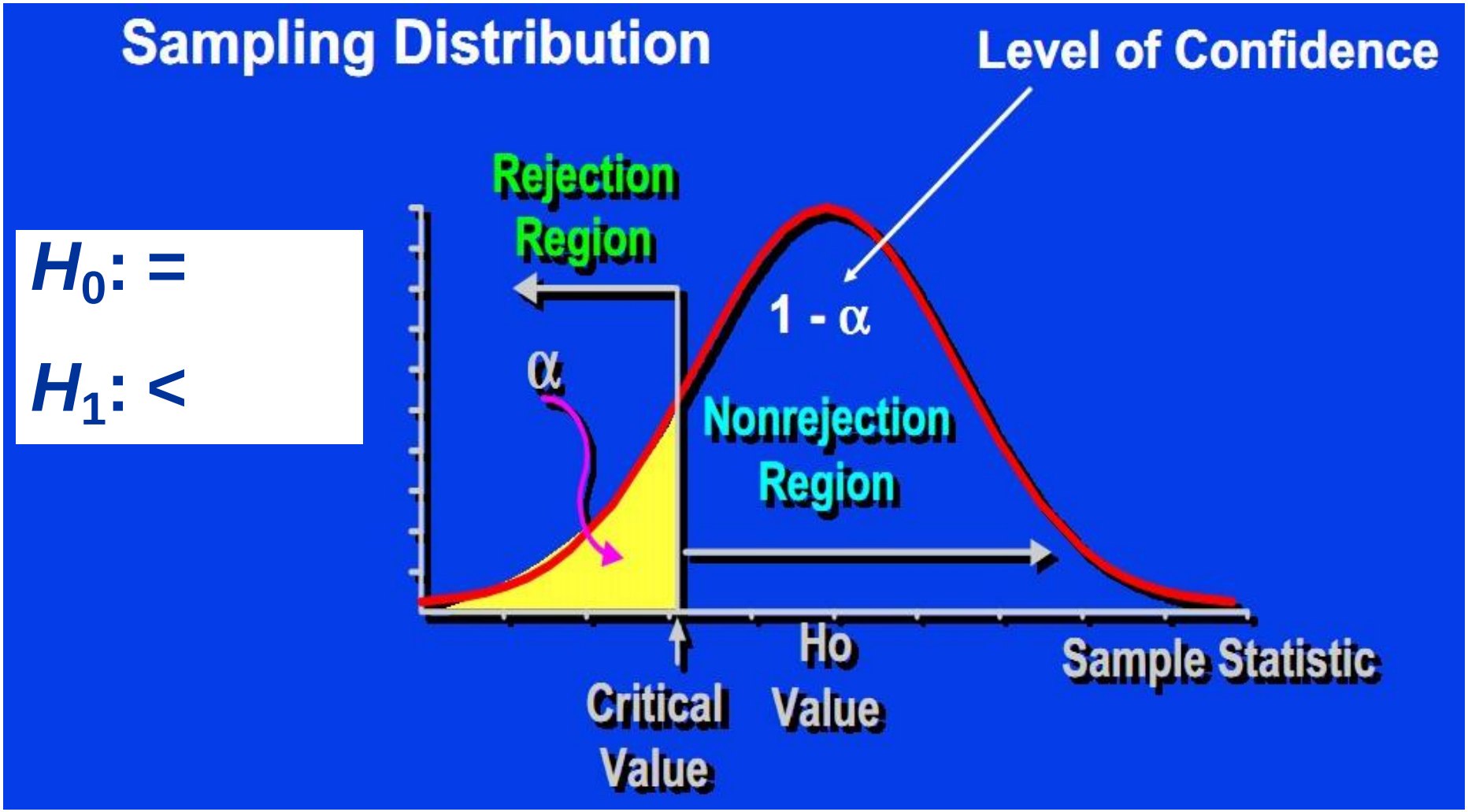
Determine the critical value

- **Find the critical value** in the distribution table for each test statistics(tabulated value) using the α - value and the degrees of freedom(where applicable)
- The critical value separates the **acceptance zone** from the **rejection zone of H_0**
- **One tailed test** – area of rejection is in either the lower or upper tail of the distribution
- **Two- tailed** – 2 areas of rejection, one in each tail of the distribution

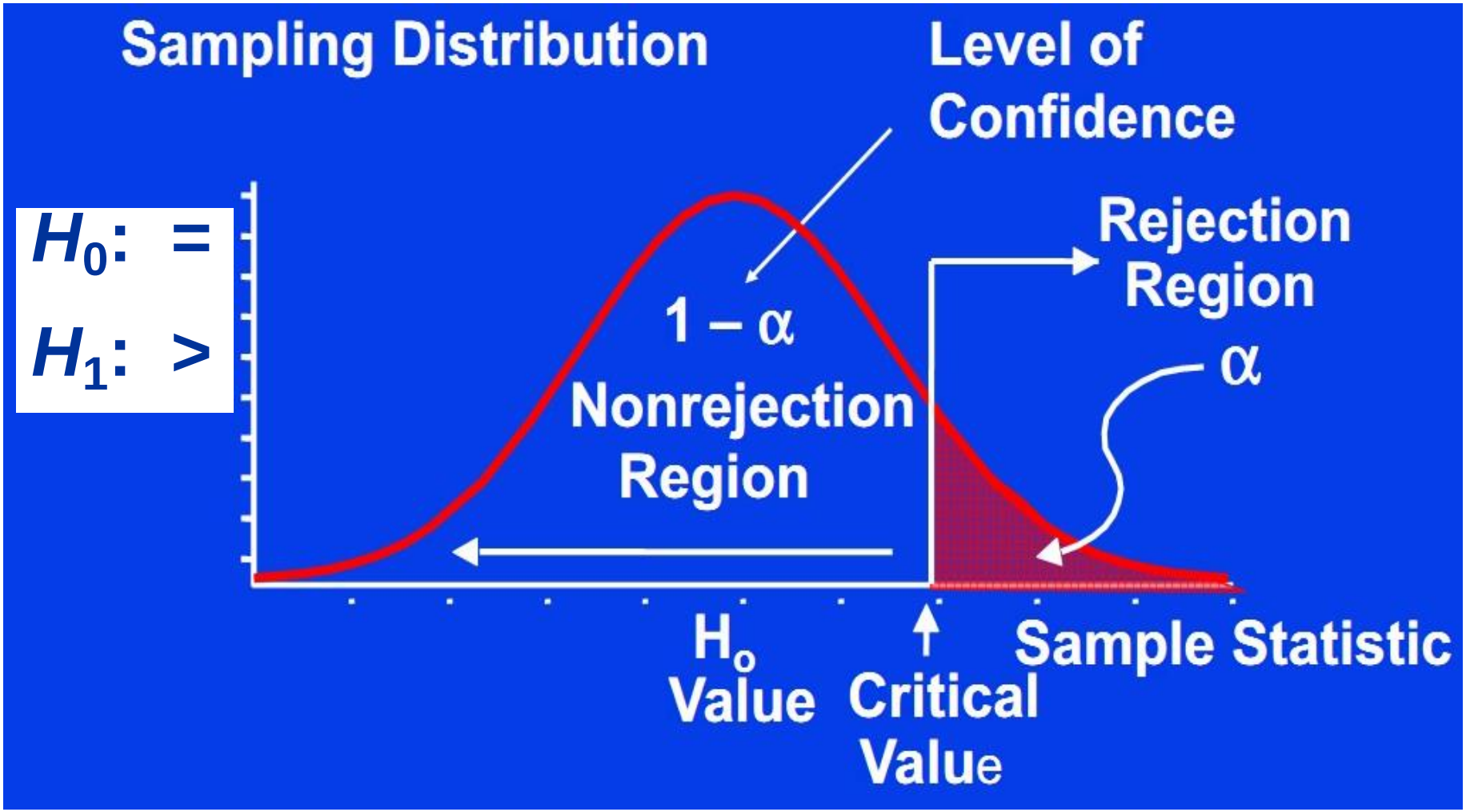
Rejection Regions (Two -tail test)



Rejection Region
(One - tail test: Left tail test)



Rejection Region
(One - tail test: Right tail test)



3) Decide on the appropriate test statistic for the hypothesis (Z, t, $\chi 2$, F, etc) and **compute the test statistic value**

Cont.

Apply the formula of the test statistics to get the calculated value

$$\text{test statistic} = \frac{\text{Estimate-Hypothesized Parameter}}{\text{Standard Error}}$$

- Compare the calculated value to the tabulated or the critical value
- Check which zone the calculated value falls into

4) Make decision as per the decision rule

There are two possible decisions:

I) Reject the null hypothesis

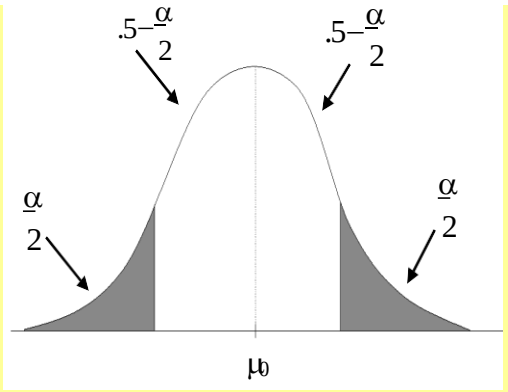
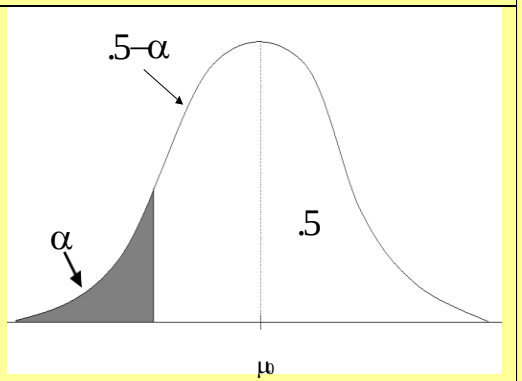
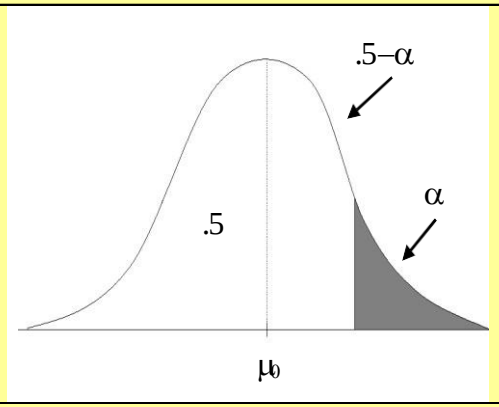
- Conclude that there is enough evidence to support the alternative hypothesis

II) Do not reject the null hypothesis

- Conclude that there is not enough evidence to support the alternative hypothesis

- If the numerical value of the test statistic **falls in the rejection region**, we reject the null hypothesis
- If the test statistic **does not fall in the rejection region**, we do not reject H_0

Decision rule

Form of H_a: $\mu \neq \mu_0$ 2-tail hypothesis If $\lvert z \rvert > \lvert z_{\alpha/2} \rvert$ Then reject the null hypothesis.	
Form of H_a: $\mu < \mu_0$ 1-tail hypothesis If $z < -z_{\alpha}$ Then reject the null hypothesis.	
Form of H_a: $\mu > \mu_0$ 1-tail hypothesis If $z > z_{\alpha}$ Then reject the null hypothesis.	

p-value

- **p-value:** is the probability of obtaining a test statistic at least as extreme ($\leq$ or $\geq$) as the observed sample value, given H_0 is true
  - Also called observed level of significance
  - Smallest value of α for which H_0 can be rejected
- **Decision :** Compare the p-value with α
  - If p-value $\geq \alpha$, do not reject H_0
  - If p-value $< \alpha$, reject H_0

Con.

- 5) **draw conclusion**

Hypothesis testing on a single population mean

- Zcal = $\frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

EXAMPLE-1:

A researcher claims that the mean of the IQ for 64 patients is 110 and the expected value for all population is 100 with standard deviation of 10. Test the hypothesis that the population mean is different from 100(α=5%).

Solution:

1. Ho:μ=100 VS HA:$\mu \neq 100$
2. Assume α=0.05, z-critical at 0,025 is equal to 1.96.
3. z=(110-100)8/10=8
4. $8 \geq 1.96$ There fore, the decision is reject the null hypothesis
5. We have sufficient evidence to support that the mean IQ of patients is different from 100.

Example-2 :

Researchers are interested in the mean level of some enzyme in a certain population. They are asking: can we conclude that the mean enzyme level in this population is different from 25? They collect a sample of size 10 from a normally distributed population with a known variance, $\sigma 2 = 45$. The calculated sample mean is = 22.

Step 1: H0: $\mu = 25$

HA: $\mu \neq 25$

Step 2: $\alpha = 0.05$, z-critical at 0,025 is equal to 1.96.

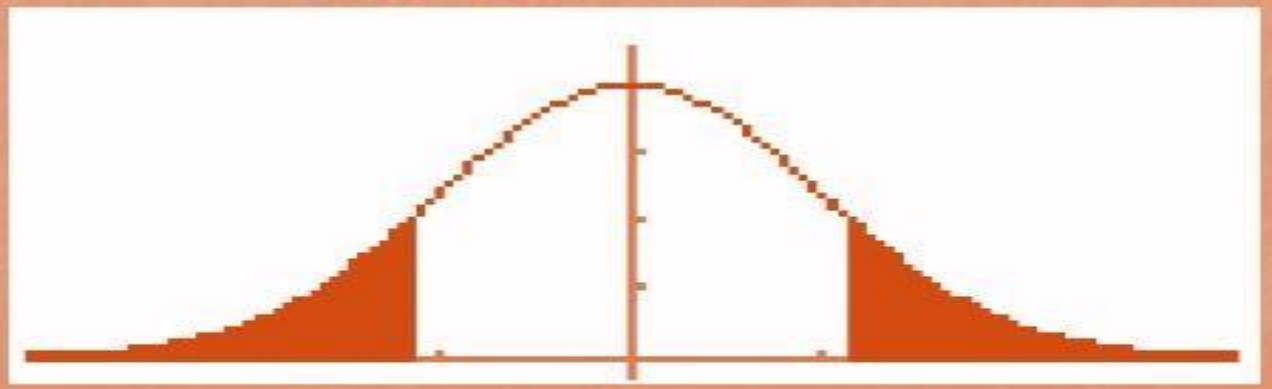
Step3:Testing a hypothesis about population mean.

The population is normally distributed. Population variance is known

$$z = (\bar{x} - \mu_0) / (\sigma / \sqrt{n})$$

- Step 6: $\bar{X} = 22$, $\mu_0 = 25$, $\sigma = \sqrt{45}$, $n = 10$

$$\Rightarrow \quad Z = \frac{22 - 25}{\sqrt{45 / 10}} = -1.41$$



Step 5: Since -1.41 falls in the acceptance region we accept the null hypothesis.

- The mean enzyme level in the population is not different from 25.

Hypothesis testing about differences between population means (normally distributed)

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- $\alpha = 0.05 \Rightarrow Z_{0.025} = 1.96$

Example:

- In a large hospital for the treatment of : mentally retarded, a sample of 12 individuals with mongolism yielded a mean serum uric acid value of = 4.5 mg/100 ml. In a general hospital a sample of 15 normal individuals of the same age and sex were found to have a mean value of =3.4. If it is reasonable to assume that the two populations of values are **normally distributed** with variance equal to 1, do these data provide sufficient evidence to indicate a difference in mean serum uric acid levels between normal individuals and individuals with mongolism? (α=5%)

Solution

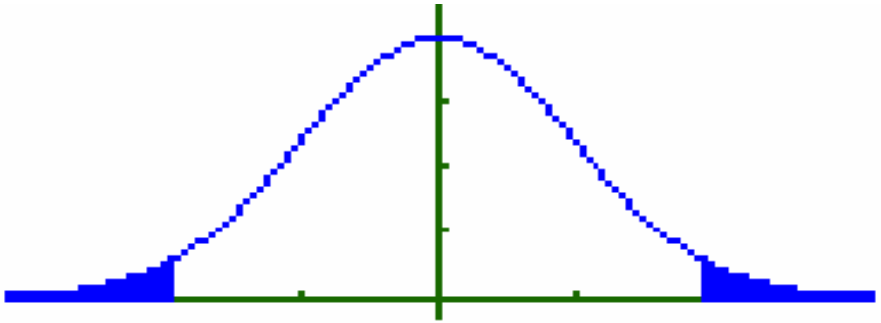
- $H_0: \mu_1 = \mu_2$
  $H_1: \mu_1 \neq \mu_2$

- $\overline{x}_1 = 4.5$ $\overline{x}_2 = 3.4$
  $n_1 = 12$ $n_2 = 15$
  $\sigma_1 = 1$ $\sigma_2 = 1$
- Z - test is the appropriate one

Solution

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- $\alpha = 0.05 \Rightarrow Z_{0.025} = 1.96$



Solution

$$Z = \frac{(4.5 - 3.4) - 0}{\sqrt{1/12 + 1/15}} = 2.82$$

- Since 2.82 > 1.96 reject the null hypothesis, i.e., there is an indication that the means are not equal.

Hypothesis testing about a single population proportion

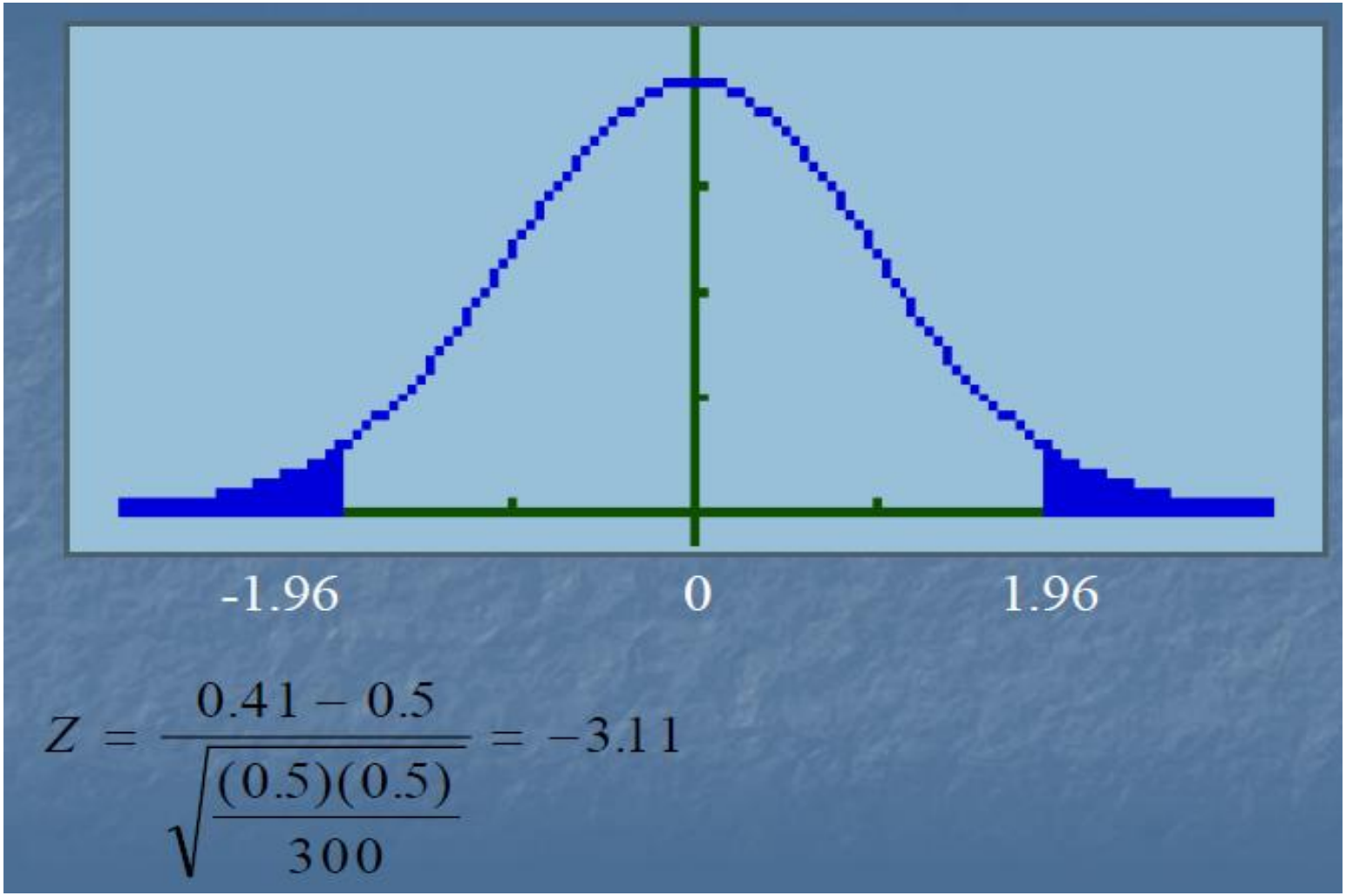
- **Example** : A survey was conducted to study the dental health practices, and attitudes of a certain urban adult population. Of 300 adults interviewed, 123 said that they regularly had a dental check up twice a year. Can we conclude that the population proportion $\pi = 0.5$? (α=5%)

Solution

- $H_0: \pi = 0.5$

  $H_1: \pi \neq 0.5$
- $P = 123/300 = 0.41$
- Test statistic, $Z = \dfrac{P - \pi_0}{\sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}}$
- $\alpha = 0.05, \quad Z_{0.025} = 1.96$

Solution



$$Z = \frac{0.41 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{300}}} = -3.11$$

Solution

- Since -3.11 < -1.96, we reject H_0. We conclude that not 50% of the population regularly have a dental check up twice a year.

Hypothesis testing for two population proportion

- The standard error of the difference is given by:

- The test statistic becomes $Z = \dfrac{(P_1 - P_2) - (\pi_1 - \pi_2)}{SE_{(P_1 - P_2)}}$

$$SE_{(P_1 - P_2)} = \sqrt{\frac{P(1-P)}{n_1} + \frac{P(1-P)}{n_2}} = \sqrt{P(1-P)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$$

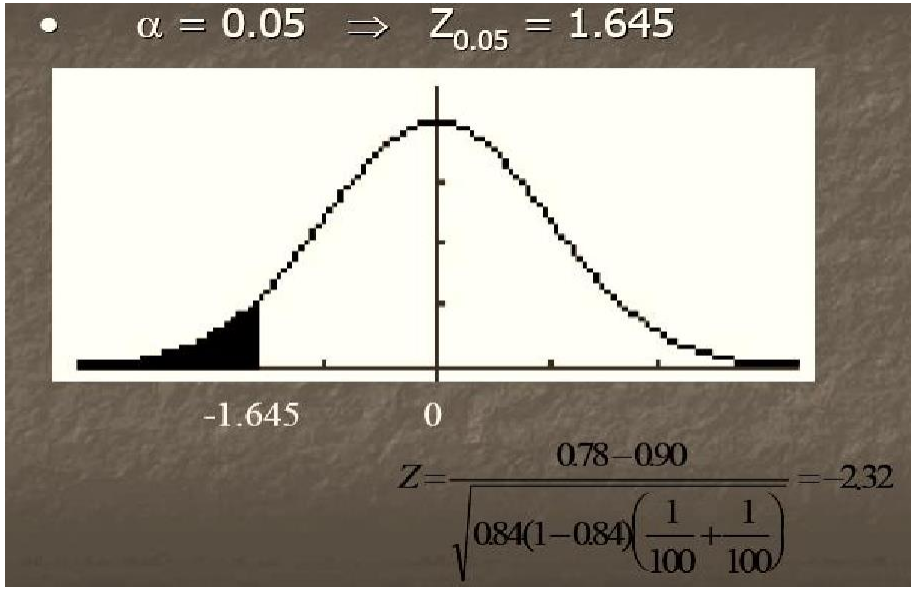
Example

- Two hundred patients suffering from a certain disease were randomly divided into two equal groups.
- Of the first group, 78 recovered within three days. Out of the other 100, who were treated by a new method, 90 recovered within three days.
- The physician wished to know whether the data provide sufficient evidence to indicate that the new treatment is more effective than the standard assuming that they have common population proportion. (α=5%)

Solution

- H0 : $\pi 1 = \pi 2$
- H1: $\pi 1 < \pi 2$
- P1 = 78/100 = 0.78, P2 = 90/100 = 0.90

$$P = \frac{100(0.78) + 100(0.90)}{100 + 100} = 0.84$$



- Since -2.32 < -1.645 it falls in the rejection region, we reject the null hypothesis.

- The data suggests that the new treatment is more effective than the standard.

Errors in making Decision

1. **Type I Error**
   - Probability of rejecting null hypothesis when its true
   - Probability of accepting a false alternative hypothesis
   - Probability of Type I Error is α (Alpha)
     - Called **level of significance/confidence**

2. **Type II Error**
   - Probability of accepting null hypothesis when its false
   - Probability of failing to reject null hypothesis when its false
   - Probability of rejecting a true alternative hypothesis
   - Probability of Type II Error is β (Beta)

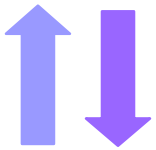
Power of a statistical test

- **Power :** The probability of rejecting the null hypothesis when it is false.
- Power = 1- β.

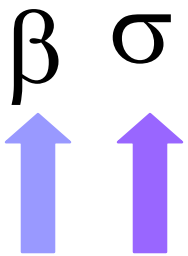
- Clearly if the test maximizes power, it minimizes the probability of Type 2 error β.

Factors Affecting Type II Error

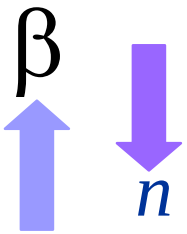
- **Significance level**
  - β Increases when α decreases



- **Population standard deviation**
  - β Increases when σ increases



- **Sample size**
  - β Increases when n decreases



Decision Results

H_0: Innocent

Jury Trial		
Verdict	Actual Situation	
	Innocent	Guilty
Innocent	Correct	Error
Guilty	Error	Correct

H_0 Test		
Decision	Actual Situation	
	H_0 True	H_0 False
Accept H_0	$1 - \alpha$	Type II Error (β)
Reject H_0	Type I Error (α)	Power ($1 - \beta$)

True positive
You're pregnant

True negative
You're not pregnant

False positive
(Type I error)
You're pregnant

False negative
(Type II error)
You're not pregnant

Exercise

1) The mean serum cholesterol level in a certain population of normal healthy men is 245 mg/dl and the standard deviation is 45 mg/dl. A clinical researcher interested in comparing cholesterol levels in this healthy population with those in men with coronary artery disease measured serum cholesterol level in a random sample of 100men who had undergone coronary bypass surgery during the preceding two year period. The mean serum cholesterol level for sampling was 267 mg/dl. can the researcher conclude that the mean serum cholesterol of men undergoing coronary bypass surgery differs from that of healthy men? Alpha=5%

Exercise

- A study was to investigate the oral status of a group of patients diagnosed with thalassemia major (TM). One of the outcome measures was the decayed, missing, and filled teeth index (DMFT). In a sample of 18 patients the mean DMFT index value was 11.3 with a standard deviation of 6.3. Is this sufficient evidence to allow us to conclude that the mean DMFT index is greater than 9.0 in a population of similar subjects? Let alpha=5%

Exercise

- A health officer is trying to study the malaria situation of Ethiopia. From the records of seasonal blood survey (SBS) results he came to understand that the proportion of people having malaria in Ethiopia was 3.8% in 1978 (Eth. Cal). The size of the sample considered was 15000. He also realised that during the year that followed (1979), blood samples were taken from 10,000 randomly selected persons. The result of the 1979 seasonal blood survey showed that 200 persons were positive for malaria. Help the health officer in testing the hypothesis that the malaria situation of 1979 did not show any significant difference from that of 1978 (take the level of significance, $\alpha = 0.01$).

Class end